

# Responsive Change: Agency Output Response to Reputational Threats\*

Moshe Maor  
Wolfson Family Professor of Public Administration  
Department of Political Science  
The Hebrew University of Jerusalem  
Mount Scopus, Jerusalem 91905  
[moshe.maor@mail.huji.ac.il](mailto:moshe.maor@mail.huji.ac.il)

Raanan Sulitzeanu-Kenan  
Senior Lecturer  
Department of Political Science and Federmann School of Public Policy  
The Hebrew University of Jerusalem  
Mount Scopus, Jerusalem 91905  
[rasulke@mssc.huji.ac.il](mailto:rasulke@mssc.huji.ac.il)

**Abstract.** How do reputational threats affect agency outputs? We undertake quantitative and qualitative analyses of reputation and outputs data regarding the fight against welfare fraud by the main service delivery agency for the Australian Government in the field of social policy. We find that an agency's response to reputational threats is endogenously differential both within the set of agency outputs and between agency outputs and other activities (a pattern we termed *responsive change*). For the former, we find that when an agency output is below average, negative media coverage leads to an increase in output in the following year. However, this relationship is nullified for agency outputs that are about average, and is reversed for outputs that are above average, i.e., these outputs tend to decrease following negative media coverage. For the latter, we find that when a reputational threat is joined by a general above-average level of outputs, the agency's drive for change is likely to be channeled into activities other than the number of units of service delivered (e.g., public relations, community engagements, stakeholder consultations, etc.).

---

\* Earlier versions of the paper were presented at the 2014 ECPR General Conference, Glasgow; the 2014 Accountability Workshop, LSE; the 2014 IRSPM Annual Conference, Ottawa, and the 2014 SOG/IPSA Conference, Jerusalem. The authors thank the *Levi Eshkol Institute for Economic, Social, and Political Research* at the Hebrew University of Jerusalem, and the *Fritz Thyssen Stiftung* (Az 10.11.2.132) for their financial support. The theoretical development of this paper was also facilitated by a grant from the *Israel Science Foundation*. The authors also thank the Australian Government Department of Human Services for its assistance in this project, as well as Alon Burstein for his superb research assistance on this paper.

How do reputational threats affect agency outputs? *Agency output* is a performance measure which refers to the number of units of goods and services produced by an agency, an increase of which could be credited to the work of the agency (Boyne, Entwistle, and Ashworth, 2010; Dunleavy and Carrera 2013). Ideally, an agency may employ a mix of input, output, process and outcome measures, and manage its activities towards achieving results, thereby measuring its performance by using outcomes (i.e., what it wishes to achieve) and impacts (i.e., what it wishes to change). However, there are situations wherein an agency may prefer to use its outputs as the main performance measure; for example, when outcomes are inherently difficult to measure; when an agency is one of many contributors to the desired outcome; when the policy domain relates to deterrence or prevention of specific behaviors, especially when the actual scope of the policy problem is unknown; when policy outcomes may not be achieved for many years, and when the purpose of the program is administrative or process oriented.

Studies that aim at identifying factors that account for agency performance (e.g., Boyne, 2003; Ashworth, Boyne, and Entwistle, 2010) have devoted limited attention to the impact of reputational considerations. However, in a recent assessment of the effects of organizational resources on agency performance, Lee and Whitford (2010) have found that an agency's public reputation, among others, has a positive impact on the agency's performance, measured by the level of goal achievement. Relatedly, Krause and Douglas (2005) found that reputational maintenance is more important to an agency than succumbing to political pressures and therefore, there is no observable relationship between political insulation and bureaucratic performance. Instead, agencies employ imitative practices in order to avoid being labeled inferior. These findings were

corroborated in a related study which suggests that differences in agency design at the federal level do not explain differences in the quality of administrative performance (Krause and Douglas, 2006). Furthermore, Carpenter has argued that audience judgments of an agency may simultaneously relate to its output performance, the expertise of its staff, its values and/or the legality of its processes (Carpenter, 2010). While these studies are insightful, important unanswered questions remain regarding the processes and mechanisms by which reputation maintenance affects agency performance. The results of this study offer a nuanced response to this question, by identifying the conditions that determine the different agency responses — *insofar as its outputs are concerned* — to reputational threats.

We analyze the relationship between an agency's reputational threats, as reflected in the printed media, and ten organizational outputs over a period of ten years. We do not assess agency reputation as a resource (Russo and Fouts, 1997), as Lee and Whitford (2012: 15) did, but rather focus on the agency's response to reputational threats through changes in its organizational outputs. The notion of reputational threat, we believe, is more nuanced than the idea of "managing the external environment" (e.g., Hecl, 1977; Moore, 1995) by relying on public managers' "political skills", often mentioned by public management scholars (e.g., Agranoff and McGuire, 2003; Moynihan and Pandey, 2005). Along with recent recognition that "[...] public organizations competed in political reputation markets for the attention and support of the public and politicians" (Lee and Whitford, 2012: 3), we think that the focus on reputational threat can provide meaningful insights into agency response. Specifically, if the concept of 'threat' to agency reputation is so crucial, does negative coverage of the agency, whether conveying

punishment for past actions or threat as guiding future conduct, actually lead to changes in agency outputs? This question is at the heart of our study.

The main empirical analysis relies on a panel study of ten different outputs produced by *Centerlink* — the main service delivery agency for the Australian government in the field of social policy and administration (Halligan, 2008) — as part of its fight against social benefit fraud, over the period 2000-2010. The fight against social benefit fraud is a classic example of the policy areas discussed earlier wherein agency output is perhaps the most important performance measure because the scope of the problem is unclear and consequently, policy outcomes and impacts are difficult to measure. Our main independent variable is the valence of the media coverage of the agency's fight against social benefit fraud over this period. Media coverage is gathered through quantitative content analysis of one national and four regional newspapers with the highest circulation in four of the Australian states and territories. Selection of agency outputs was based on a quantitative analysis of survey responses by eighteen senior executives in charge of the fight against welfare fraud at the agency, regarding the visibility, intrusiveness and chances of failure (i.e., risk) of policy tools employed in this fight. Interpretation of the findings was further assisted by eighteen semi-structured interviews with senior executives employed at *Centerlink* and at the Department of Human Services.

We find that an agency's response to reputational threats is endogenously differential both within the set of agency outputs and between agency outputs and other activities. For the former, we find that when an agency output is below average, negative media coverage leads to an increase in that output in the following year. However, this

relationship is nullified for agency outputs that are about average, and is reversed for outputs that are above average, i.e., these outputs tend to decrease following negative media coverage. Since negative media coverage affects the agency's motivation to undergo changes in its performance, the magnitude of changes in this pattern is minimal following positive coverage, and highest following negative coverage. We call this pattern *responsive change*, which reflects greater volatility in agency activities when the organization is faced with a reputational threat and low volatility when its reputation is not threatened. This endogenous process of agency evaluation of its response to reputational threat – which relies, among others, on data regarding previous year outputs – implies distinct relationships between organizational memory (Walsh and Ungson, 1991; Weinberger, Te'eni, and Frank, 2008), reputation management and agency outputs. Regarding an agency's differential response between agency outputs and other activities, we find that when a reputational threat is joined by a general above-average level of outputs, the agency's drive for change is likely to be channeled to activities other than the number of units of service delivered, e.g., public relations during natural disasters, community engagements, stakeholder consultations, and so on.

We seek to make three contributions to public management scholarship. First, we intend to link the literatures on the effects of bureaucratic reputation and those dealing with the determinants of agency performance. These two literatures have developed independently despite the fact that agency performance may be measured by “organizational outputs” (Boyne, Entwistle, and Ashworth, 2010: 4), which can be modified following a managerial decision, and can therefore provide insight into an agency's reaction to a reputational threat. Second, despite the acknowledged importance

of public reputation (Lee and Whitford, 2010) as well as the agency's need to identify and neutralize threats originating in its external environment (Cho and Ringquist, 2010: 59), we can find little research that examines the effect of reputational concerns on organizational outputs in the public sector. We intend to fill this gap by presenting and testing a framework that is useful for research on the potential role of reputational concerns in determining agency outputs and hopefully also for various other performance measures. This research therefore contributes to the growing body of literature that seeks to render much more specific and micro-calibrated measurements and associations of reputation with other variables, not least, agency behavior. Third, and most importantly, attention to reputational concerns enables us to identify the endogenous nature of the agency's output response, one that does not depend on regulation, structures, client characteristics, or on the occurrence of some exogenous shock. Instead, the process of differentially responding to negative media coverage (Gilad, Maor, and Ben-Nun Bloom, 2013), given the level of recent agency outputs stored in the agency's memory and retrieved when necessary, focuses our attention on those considerations of the agency itself that make changes in outputs possible. This is perhaps the most important comparative advantage of employing a reputational framework. The remainder of this paper unfolds as follows. In the second section, we present the analytical framework; in the third we describe the methodology; in the fourth, the empirical findings; and in the fifth, we discuss the theoretical implications of our findings.

## **Analytical Framework**

### *Reputation and Agency Responsiveness*

The notion of reputational risk is one of the cornerstones of reputation-driven conceptual frameworks in political science nowadays (e.g., Carpenter, 2001, 2002, 2004, 2010a; Carpenter and Krause, 2012; Gilad and Yogev, 2012; Krause and Douglas, 2005; Krause and Corder, 2007; Maor, 2007, 2010, 2011; 2014, in press; Moffitt, 2010; Moynihan, 2012; Whitford, 2002, 2003). At the outset, bureaucratic reputation is defined as a set of symbolic beliefs about the unique or separable capacities, intentions, roles, obligations, history, and mission of an organization, that are embedded in a network of multiple audiences (Carpenter, 2010a: 33, 45). This definition centers on the evaluation of the organization's unique character and activities by multiple audiences. "Reputation uniqueness", according to Carpenter (2001: 5), refers to the demonstration by agencies that they can create solutions (e.g., expertise, efficiency) and provide services (e.g., moral protection) found nowhere else in the polity.

Recent findings have related to the consequences of reputational concerns for the way agencies approve some drugs more quickly than others (Carpenter, 2002) and allocate resources across tasks (Gilad, 2012), for endogenous construction of jurisdictions (Maor, 2010) and observability of decisions and errors (Maor, 2011), and for the duration of enforcement decisions (Maor and Sulitzeanu-Kenan, 2013). Other studies have demonstrated the extent and the ways regulatory agencies manage their reputations through the strategic use of communication (Maor, Gilad, and Ben-Nun Bloom, 2013; Gilad, Maor, and Ben-Nun Bloom, 2013). Carpenter (2001, 2010a) has offered some generalized answers to issues regarding reputation and power, and has summed up the contribution of this scholarly literature: "The lesson of this scholarship is that, when

trying to account for a regulator's behavior, *look at the audience*, and *look at the threats*" (Carpenter 2010b: 832; *italics in original*).

Taking this lesson on board, Maor and Sulitzeanu-Kenan (2013) analyzed the impact that reputational concerns have had on 'time-to-enforcement' decisions. Focusing on two enforcement divisions within the FDA's Center of Drug Evaluation and Research, they found that, as media coverage of the FDA's consumer protection responsibilities becomes more positive, the agency takes enforcement decisions more slowly; in contrast, more critical media coverage leads to quicker action by the FDA. The current study attempts to take a further step, and assess the effect of such concerns more directly on the *quantities* of goods and services produced by an agency. Assuming that an agency safeguards its unique reputation (e.g., Carpenter, 2001, 2010a; Heimann, 1997; Krause and Douglas, 2005; Maor, 2010, 2011; Quirk, 1980), it is reasonable to expect that agency outputs will tend to increase under media coverage that is negative in tone, as compared to media coverage with a positive tone. This is because such an agency will have an incentive to compensate for the reputational loss derived from the negative media coverage by increasing agency outputs. Although this is not an intuition that is necessarily compelling, as it would seem to matter what the specific output is, as well as what the negative coverage is specifically about, it provides an initial theoretical stepping stone suggesting that the agency under attack is likely to demonstrate visible and easily understandable evidence for an increase in its outputs. Thus, we hypothesize that:

*Hypothesis I: When faced with negative media coverage, an agency will increase its level of outputs*



### *The Notion of Selective and Differential Responsiveness*

The second step taken in this study is to empirically test a more nuanced theory regarding the effect of reputational threats on agency outputs. We propose that this effect is moderated by the agency's own judgment of its current outputs. While organizations are expected to be responsive to reputational threats, their response is plausibly the product of integrating such external signals with internal judgments regarding their relevance and validity, given other types of information available to the agency or stored in its organizational memory to be retrieved when necessary (Walsh and Ungson, 1991; Weinberger, Te'eni, and Frank, 2008). This integration is likely to moderate the response of the agency to reputational threats. Maor, Gilad, and Ben-Nun Bloom (2013) have demonstrated that a regulator *selectively* responded in the media to external allegations, given its assessment of their relative threat to its reputation. "Selective response" refers to a variation of agency behavior between *levels* of a particular type of response. Specifically, they have shown that a regulatory agency tends to keep silent on issues regarding which it generally enjoys a strong reputation, and on issues that lie outside its distinct jurisdiction, while responding to opinions about core functional areas in which its reputation is weaker and areas wherein its reputation is still evolving.

If agency budget is an analogue of agency reputation, the aforementioned example ties in nicely with Macmillan's (2012) finding that the U.S. Department of Justice's Anti-Trust Division increases enforcement output when their funding is lower than they deem optimal, while engaging in preserving status-quo enforcement levels when their funding is higher than they deem optimal. This explicit tradeoff between

financial security and policy goals is undertaken by agencies in order to protect themselves from external pressure brought about by appropriations politics.

In the context of performance, we therefore expect an agency's outputs patterns to be selectively responsive to reputational threats when the agency's internal evaluation of its outputs is low, and to exhibit less responsiveness when the agency evaluates its outputs as high. We employ average agency outputs as a proxy for the agency's evaluative reference point, and recent actual relative output level as a proxy for the agency's internal evaluation, and hypothesize that:

*Hypothesis II: The effect of negative media coverage on an agency output level will be moderated by the recent level of that output, i.e., the effect will be stronger following periods of below average levels, and weaker following periods of above average levels.*

As our model of agency response to reputational threats is temporally dynamic, we must assume that 'regression to the mean' (RTM) plays a role in this process, unless shown otherwise (Barnett et al. 2005). This assumption is particularly relevant as our second hypothesis predicts that recent (lagged) agency outputs are expected to influence the level of consequent outputs. Taking regression to the mean as our baseline expectation (rather than stability), and drawing on Macmillan (2012) and Maor and Sulitzeanu-Kenan (2013), our second hypothesis entails greater stability (weaker regression to the mean) under positive media coverage, and greater responsiveness (stronger regression to the mean) under negative media coverage.

Another potential intervening factor that may also condition the agency's capacity to initiate change in general, and in response to external triggers in particular, is

organizational slack. Organizational slack is “the pool of resources in an organization that is in excess of the minimum necessary to produce a given level of organizational output” (Nohria and Gulati, 1996). If this is the case, the hypothesized observed relationship between agency responsiveness and recent output levels can actually stem from the availability of slack, rather than the evaluation by the agency of its outputs. In our empirical analysis we include a measure of slack in order to account for this potential alternative process.

Ultimately, we hypothesize that a public agency’s response to reputational threats will be undertaken, among others, through changes in its organizational outputs. Such a finding entails that reputational concerns create a response mechanism that produces institutional outcomes (agency outputs) over and above agency slack, the content of rules, guidance, procedures, structures, and office holders. Our attention now turns to the methodological section.

## **Research Design**

We investigate the effect that reputational considerations have on agency outputs by focusing on a single case which is theoretically illuminating, relying on two novel datasets that have been constructed as part of our research, and contain agency outputs and media coverage over a decade. These datasets allow us to estimate an agency’s strategic response to reputational threats by analyzing systematic changes in its organizational outputs in a panel design.

### *The Policy Domain*

The policy domain examined here is the prevention and investigation of social benefit fraud. Welfare fraud occurs when individuals or organizations acquire social benefits from a government agency or department, either by obtaining payments to which they are not entitled or by withholding funds that they should pay over to the agency or department. This policy domain was selected for three reasons, which also contributed to the increased attention devoted to the issue by the general public, the media and policymakers during the last decade. First, in recent years there has been growing recognition of the negative impact that welfare fraud has on the economies of industrialized countries that acknowledge public responsibility for the provision of some social support and spend very substantial amounts on benefit payments (e.g., UK National Audit Office, 2006; UK Treasury & National Audit Office, 2008; UK Revenue & Customs and Department for Work and Pensions, 2010). Second, criminal prosecutions as well as other anti-fraud measures have created an ongoing controversy in welfare countries regarding the legitimacy of the welfare system, the balance between a punitive approach and prevention of fraud before it occurs, and the stigmatization of welfare recipients (e.g., Bradbury, 1988: 26; Prenzler, 2011). Third, the fight against welfare fraud has gained tremendous momentum following the 2008 financial meltdown and the consequent economic crisis. Many industrialized countries have made combating fraud a major issue in their government objectives, expanding existing prevention and detection measures, and introducing new ones (Prenzler, 2011).

Despite the growing recognition in recent years of the damaging impact of welfare fraud, most scholars do not place the fight against social benefits fraud at the center of analysis. There are some notable exceptions: Etzioni (1982) has investigated

anti-fraud initiatives based on prevention that is grounded in an understanding of the constituencies involved; Grover (2005) has looked at advertising social security fraud, and Rowlingson et al. (1997) have studied the role of penalties in this context. Moreover, McKeevers (1999a) has examined public reaction to benefit fraud in the U.K., and the rationale for legislative changes in the fight against benefit fraud (McKeevers 1999b). Walsh and Martson (2010) have studied the conflation of social security fraud and overpayment. Sandberg (1993) has shown that tolerance for different forms of abuse has generally decreased in Sweden. And Chunn and Gavigan (2004, 2006) have analyzed “welfare cheats,” the place this concept occupies in attacks on the poor, and the theoretical and empirical questions related to regulation, control, and the relationship between them at particular historical moments. These studies have provided descriptive and normative accounts of this policy domain. However, the possible impact reputational concerns may have had on the fight against welfare fraud has remained uncharted territory.

#### *Australia’s Centrelink Agency*

Australia has provided welfare benefits to vulnerable and disadvantaged citizens since 1909, and during the period under examination (i.e., 2000-2010), almost one-third of Australians received some kind of direct welfare payment (Prenzler, 2011). In 2008-09, for example, Australia’s federal welfare agency *Centrelink* distributed approximately \$75b to 6.8 million recipients across 140 benefit types on behalf of 27 government departments and agencies (*Centerlink*, 2009: 28; Figures are in U.S. dollars). The availability of numerous benefits during the period under investigation has created

opportunities for minor and major welfare fraud. Not surprisingly, media coverage of fraud has revolved around major cases of welfare fraud, alerting the public to the anti-fraud work of government agencies and departments and to the costs of fraud (Prenzler, 2011: 4). In addition, media coverage has fuelled popular opinion against welfare cheats (Prenzler, 2011: 2). Still, estimates of the level of fraud have ranged from “[...] a significant number of people [...]” (Bajada, 2005: 184), to less than half of one percent of social security debt (Karvelas, 2008: 4). Beyond estimates, although conviction rates during the period under investigation stood at around 3000 per year, representing 0.04% of *Centerlink*’s customers, the derived losses of these cases involved approximately \$105m per year in gross savings and amounts targeted for recovery (Prenzler, 2012: xii).

The agency under investigation is *Centerlink*, which was established in 1997 as the main service delivery agency for the Australian government in the field of social policy and administration (Halligan, 2008). In October 2004, *Centerlink* autonomy was considerably curtailed with the establishment of the Department of Human Services, which has been entrusted with the development of service delivery policy and the provision of access to social, health and other payments and services (Halligan, 2007). Consequently, the agency’s CEO who had served since 1996 resigned, and in 2005-06, new governance arrangements were created. The agency was formally incorporated into the Department of Human Services in 2011 as a result of the *Human Services Act 2011*, with the department retaining the *Centerlink* brand name as part of its set of main policy programs.<sup>1</sup> The period under examination covers 2000-2010, thus encompassing most of the period in which *Centerlink* was an autonomous and semi-autonomous agency.

### *Public Agency Outputs*

Measuring outputs in the public sector is notoriously more complicated than doing so in the private sector, where a price-weighted measure of overall outputs can be relatively easily calculated (Dunleavy and Carrera, 2013). In contrast to the private sector, some public goods and services are provided for free or at a subsidized price, some are mandatory, and others are imposed on citizens (e.g., prisons). Additionally, ethical considerations bar public institutions from targeting particular segments of the population (e.g., high income people), and in many policy domains it is not practical to rely on citizens' satisfaction as a general measure of outputs and performance (e.g., nuclear policy). This complexity is deepened when no cross-agency comparison is forthcoming due to the fact that only one agency produces the relevant goods or services in a given polity. For these reasons, the recommended practice of measuring public outputs relies on internal measures of social delivery. Such measures need to encompass the various types of outputs for lack of a common currency of measure (Kelly 2005). In line with these considerations, the measure of agency output adopted here relies on the entire set of agency activities, which were identified by senior agency officials interviewed, as critical for achieving the agency's overall task of fighting welfare fraud. This approach requires two assumptions that should be made explicit. First, that at the level of individual output, the units of goods and services have a relatively homogeneous quality. And second, that each of the outputs under consideration provides a comparable contribution to the task performed by the agency.

*Centerlink's* outputs data over the period under investigation have been obtained from both the agency's annual reports and the agency itself upon our request. Table 1

presents the list of agency outputs which were selected. We focus on the main detection strategies, namely, eligibility reviews, service profiling on the basis of recipients' characteristics (i.e., a risk-based approach), job capacity assessment and data matching with government and non-government institutions (e.g., universities). We omit the reporting of suspected fraud by the general public because the agency, once establishing the infrastructure (e.g., fraud tip-off line), has played a passive role in employing this strategy for detection purposes. In addition, we focus on the aggregate measure of fraud investigations. This measure consists of investigations which vary by target (e.g., identity-based and cash-economy investigations), source (e.g., tip-offs), mode (e.g., optical surveillance), and location (e.g., investigations of customers living in Australia or overseas). We also take into account the main agency activities which are derived from investigations, namely, prosecution referrals, formal repayment agreements, and debt recovery by private contractors. In addition, we consider the main measure employed by the agency to increase public awareness regarding the fight against welfare fraud as well as deterrence, namely, public campaigns. Figure 1 presents the temporal variation in each of the ten agency outputs during the period under examination. To allow inter-output comparability, raw measures were converted into standardized measures (standard deviations).

---

**TABLE 1 ABOUT HERE**

---

**FIGURE 1 ABOUT HERE**

---



### *Data and Operationalization*

Our measure of reputational threat rests on a quantitative content analysis of the media coverage of the agency's performance in the fight against social benefit fraud. To conduct our analysis, we made use of a national newspaper – the *Australian* – as well as four daily newspapers – *the Age*, *the Courier Mail*, *the Hobart Mercury*, and *the Sydney Morning Herald*; each is the most widely-circulated newspaper in one of four Australian states. Using the *Lexis-Nexis* database, we identified all articles published between January 1, 2000 and December 31, 2010 that addressed the agency's activities in the fight against welfare fraud.<sup>2</sup> These newspaper articles were filtered in order to exclude letters and obituaries. Furthermore, in order to avoid endogeneity of the press valence measure, the measure of media valence relied selectively on press articles in which the sources did not include the agency; thus the valence is based solely on external sources. The remaining data consisted of 771 articles which were coded for valence (positive, negative and neutral). Appendix 1 provides examples of negative and positive claims regarding the fight against welfare fraud undertaken by *Centerlink*. In order to assess the reliability of the coding, a random sample of 20 articles was coded independently by three coders. Treating the valence coding as ordinal resulted in acceptable levels of inter-coder reliability (Krippendorff's  $\alpha = 0.83$ ). Figure 2 presents the temporal variation in media valence for the period under examination. In addition, eighteen semi-structured interviews with senior executives employed at *Centerlink* and at the Department of Human Services were undertaken during January 2012 as part of this research. All interviews were recorded, fourteen of which were conducted at the agency's offices, and the remaining four over the phone. Interviewees were selected based on their executive

role covering the top three administrative ranks at *Centerlink*. After each interview, the interviewees were asked to rank the agency’s policy tools (1-10) according to their level of public visibility, intrusiveness and chances of failure.

---

**FIGURE 2 ABOUT HERE**

---

As noted above, we employ actual pervious year agency outputs as a proxy for the agency’s internal evaluation of its performance. In order to do so it is important to assess whether this proxy in fact reflects another variable – organization slack – that may account for the hypothesized relationship.<sup>3</sup> One may expect that the agency’s responsiveness to reputational threats will be moderated by organizational slack; thus, such slack is likely to be correlated with the level of output. In order to account for the level of slack in the agency we created a measure of slack by dividing the agency’s previous year standardized budget by previous year outputs.

To test our hypotheses we conducted a joint analysis of all ten agency outputs, by converting their yearly raw measures into standardized measures. This research design provides a cross-sectional (output) panel (year) dataset. To analyze the data, we estimated the following regression model:

$$\Delta Y_{ot} = \alpha + \beta_1 V_{t-1} + \beta_2 Y_{ot-1} + \beta_3 V_{t-1} Y_{ot-1} + \beta_4 \frac{\text{budget}_{t-1}}{Y_{ot-1}} + \beta_5 V_{t-1} \frac{\text{budget}_{t-1}}{Y_{ot-1}} + \beta_6 T + \epsilon_{ot}$$

where  $\Delta Y_{ot}$  represents the *change* in standardized performance of policy output  $o$  in year  $t$ ,  $V_{t-1}$  is the measure of one-year lagged press valence (a higher value indicates more positive coverage),  $Y_{ot-1}$  is the one-year lagged performance of the policy output,

$\frac{budget_{t-1}}{Y_{ot-1}}$  is the measure of one-year lagged slack,  $T$  represents the yearly trend, and  $\epsilon_{ot}$  represents stochastic error. The coefficient ( $\beta_3$ ) of the interaction term  $V_{t-1}Y_{ot-1}$  estimates the difference in the agency's response to negative media coverage over varying levels of previous year outputs, and thus provides the theoretically central estimation.

Our analysis is based on changes in a set of agency outputs over time. With such data one must assume that regression to the mean (RTM) may affect the temporal dynamic of the data. Following the method suggested by Barnett et al. (2005) we evaluate the magnitude of the RTM effect by plotting a scatterplot of change in the standardized output (follow-up minus one-year lagged output) against one-year lagged output, presented in Figure 3. The fitted line was obtained by linear regression of the change values on baseline values. A clear pattern of RTM is apparent in the plot, as outputs that follow unusually low levels have tended to increase (so that change values are likely to be above zero), outputs that follow unusually high levels have tended to decrease (so that change values are likely to be below zero), and average output levels have tended to be followed by no change (on average). This pattern depicts a classic RTM effect, estimated at -0.634 standard deviation change for a+1 standard deviation in the one-year lagged performance. Note that the RTM effect is expected to be uniform for a given level of lagged agency outputs, while our second hypothesis predicts that the change in outputs will also depend on media valence by estimating the interaction between lagged outputs and media valence; therefore it is not expected to bias our estimations.

---

**FIGURE 3 ABOUT HERE**

---

Given that all the outputs are produced by the same agency, we should expect a cross-sectional correlation (e.g., the agency's finite budget entails that an increase in one output leaves fewer resources for others) (Beck and Katz, 1995; Driscoll and Kraay, 1998). In order to estimate standard errors that are robust to disturbances resulting from heteroscedasticity, autocorrelation, and cross-sectional dependence, we used the Driscoll and Kraay (1998) approach, implemented by the `xtscc` Stata command (Hoechle 2007). Data are clustered by policy outputs.

## **Results**

Table 2 presents five regression analyses with Driscoll-Kraay standard errors, and *change* in standardized agency output as dependent variable. Model one provides a baseline model by regressing change in standardized agency outputs on one-year lagged standardized outputs and time.<sup>4</sup> Model 1 indicates that one-year lagged output and time are fairly good predictors of consequent year change in output. The former represents a clear empirical implication of RTM, and the latter, the particular overall trend in the data over the research period. This model accounts for 45% of the variance in the change in agency outputs, and provides a baseline for evaluating the contribution of the variables of theoretical importance for predicting agency output in the following models.

Model 2 adds one-year lagged media valence, in order to test the first hypothesis. This analysis indicates no significant association between press valence and agency output, thus providing no support for the first hypothesis. However, it is possible that the moderating mechanism suggested in our second hypothesis does not allow us to detect a statistically significant main effect of media valence.

In order to test our second hypothesis, Model 3 adds an interaction term between press valence and one-year lagged output. The fact that the coefficient for this interaction term is statistically significant suggests that the change in agency output predicted by previous year media valence is significantly different when lagged output is relatively high and low, providing support for our second hypothesis. Models 4 and 5 further assess the robustness of this finding by controlling for the potential effect of slack. Model 4 adds a simple control for slack and model 5 adds an interaction term between lagged media valence and slack – allowing us to estimate the fit of the alternative explanation to the data. The size and statistical significance of the ‘valence X lagged-output’ interaction term remain substantively unchanged across models 4 and 5. Note that this finding holds although model 5 provides support for the expectation that slack also moderates the agency’s output response to media valence. Interpreting the results of model 5 suggests that when lagged output level is average, media valence appears to have no effect on subsequent output level ( $\beta=.004$ ,  $p=.992$ ). When lagged output level is one standard deviation above average, media valence appears to have a positive effect on the change in output ( $\beta=.624$ ,  $p < .01$ ), and when lagged output level is one standard deviation below average, the effect of media valence is negative ( $\beta=-.413$ ,  $p < .01$ ). The difference between the latter two predictors is statistically significant ( $p < .001$ ).

---

**TABLE 2 ABOUT HERE**

---

These findings provide support for hypothesis two, while introducing another unexpected finding. As predicted by hypothesis two, when the recent level of an agency

output is relatively low, the agency appears to increase the level of that output the more negative its media coverage is, and this relationship is absent when recent output level is about average. However, when the recent level of an agency output is relatively high, the subsequent level of that output appears to decrease the more negative its media coverage has been in the preceding year. Figure 4 graphically presents these findings (estimates based on model 5). The x-axis represents three levels of one-year lagged media valence – ‘positive’ (90<sup>th</sup> percentile of the valence score), neutral (50<sup>th</sup> percentile) and negative (10<sup>th</sup> percentile). The y-axis depicts consequent change in standardized output. The three colored bars present the relationship between the two dimensions (media valence and output-level change) across three levels of lagged agency performance – ‘low’ (-1SD, blue bar), mean (red bar), and ‘high’ (+1SD, green bar), with 95% confidence intervals. While the RTM effect results in the overall vertical direction of each of different bars (in general, positive changes follow low performance, and negative changes follow positive performance), the pattern of change in the size of the bars depicts a varying effect of media valence across the levels of output. Following positive media coverage, changes in agency output levels are minimal, suggesting relatively minor improvement when output level is low and similar magnitude of decline when output level is high. In both cases the changes are smaller than the mean RTM effect. However, following negative media coverage, changes in agency outputs are largest, resulting in a strong improvement when recent output level is low and a large decline when it is high. These changes are clearly larger than the effects of positive coverage, as well as the mean RTM effect. In the following section we further discuss these findings.

---

**FIGURE 4 ABOUT HERE**

---

**Discussion**

Fifteen years ago, the topic of bureaucratic reputation was not even a promissory note, let alone a tangible reality. A decade of research makes it clear that reputational concerns influence agency behavior (for a review see Maor, in press). The findings of this study add another layer to the growing interest in this topic, suggesting the agency's responsiveness to reputational threats is not automatic, but rather dependent on the agency's internal evaluation of its performance. When agency output is below average, negative media coverage leads to an increase in that output in the following year. However, this relationship is nullified for agency outputs that are about average, and is reversed for outputs that are above average – i.e., these outputs tend to decrease following negative media coverage. While the former findings confirm hypothesis two, the latter was not expected, and an ex-post development of our theoretical framework is suggested.

The results depicted in Figure 4 indicate that *change* in the allocation of agency outputs is largest when media coverage is negative, and minimal when it is positive. This finding conforms to the notion that agencies are sensitive to reputational threats, and that this sensitivity is also manifested in their outputs. This process requires an elaboration. The main challenge faced by an agency, as any other organization for that matter, is how to allocate its limited resources to a set of outputs – an exercise that determines the particular mix of outputs the agency will undertake. In this context, the performance

responsiveness of an agency is likely to abide by the following logic: When its reputation is challenged (e.g., by negative coverage), the agency is expected to look for effective ways to respond in order to address its reputational decline. The alternatives available to the agency are given by the possible different mixes of its outputs. This process of evaluating alternatives for action and selecting the one to be pursued is undertaken by relying, among others, on performance information stored in the agency's organizational memory. Thus, when the agency faces a reputational threat, it may respond by seeking useful *changes* in its performance. This appears to be guided by the agency's own evaluation of its recent performance in each of its outputs, leading it to increase lagging ones (below average outputs), at the expense of high-performing activities (above average outputs), which therefore tend to decrease. Since negative media coverage affects the agency's motivation to undergo changes in its performance, the magnitude of change in this pattern is minimal following positive coverage and highest following negative coverage. We call this pattern *responsive change*, which reflects greater volatility in agency activities when the organization is faced with a reputational threat and low volatility when its reputation is not threatened.

This finding – a differential response to reputational threats *within* the set of agency outputs – relies on our quantitative analysis. Our qualitative analysis suggests a similar pattern in the investment of agency resources *between* outputs and other activities. Specifically, when a reputational threat is joined by a general above-average level of outputs, the agency's drive for *change* is likely to be channeled to activities other than the number of units of service delivered – for example, public relations during natural disasters, community engagements, stakeholder consultation, and so on. Assuming that



the agency's resources are finite, such allocation choices will result in a decline in agency outputs. A classic example was recorded during Tropical Cyclone Yasi in North Queensland and bushfires in Western Australia in early 2011. *Centerlink* mobilized 2/3 of its staff to recovery centers in the disaster area,<sup>5</sup> resulting in a 72% fall in prosecution referrals in that year (Centerlink 2010-11, 151).<sup>6</sup> This investment in public relations during a natural disaster was designed as an attempt to “demonstrate [... the agency's] service delivery capabilities at its finest” (Centerlink 2010-11, 4). *Centrelink* set up mini-mobile offices; its employees were provided with green shirts with the agency's logo, and the agency used banners on its website in order to build a profile around the presence of its employees on the ground. At the same time, the agency's communications team gave around 600 media interviews to reinforce the message and the positive attributes of what the agency was doing. The message was that if one needs help, one is welcome to the agency's recovery centers in the area affected, but one is warned that the agency will do the background checks. So the agency communicated that it was helping those in need but at the same time ensuring that its integrity remained intact.<sup>7</sup> These efforts to advance reputation for effective delivery of crisis payments combined with fraud prevention have been evident since 2010.<sup>8</sup>

The findings drawn from the quantitative and qualitative analyses suggest that an agency's response to reputational threats is endogenously differential both within the set of agency outputs, and between agency outputs and other activities. In other words, an agency is responsive to reputational threats, but this responsiveness is not endogenously *selective*, but rather endogenously *differential* – i.e., the response varies between *types* of actions (e.g., between different outputs or between outputs and other activities),

depending on its self-evaluation, rather than between *levels* of a particular type of response. These findings conform to recent findings regarding the effects of reputational threats on the duration of enforcement decisions in another and very different agency (Maor and Sulitzeanu-Kenan, 2013), extending the notion of endogenously selective responsiveness (Maor, Gilad, and Ben-Nun Bloom, 2013), and the above findings provide another example of an endogenously differential response (Gilad, Maor, and Ben-Nun Bloom, 2013). These findings also empirically demonstrate distinct relationships between organizational memory, reputation management and performance.

Additional research with other performance measures is required in order to demarcate the extent to which these findings are general. Suitable measures may include judgments of internal or external stakeholders regarding performance (Walker and Boyne, 2006); organizational inputs and/or outcomes (Cohen and Eimicke, 2008; Linder, 2004), and qualitative performance data (Blasi, 2002). Further unexplored questions revolve around whether changes in public agency outputs have been undertaken across the board, or whether they reflect a strategic decision limited to a carefully selected set of agency outputs. Further research is required regarding the relationship between an agency's reputation and agency behavior over longer time periods as some variation in media coverage may be historically contingent, that is, coming from a short period of time of a decade or so. In addition, this research has dealt solely with one agency, although the findings appear to conform to previous studies on very different agencies. Further research should be devoted to enhancing our understanding of how the reputational considerations of one agency may differ across policy domains and national settings. Although much work needs to be done to unravel how an agency's basis of

reputation impacts upon its behavior, we hope that this study has modestly provided a contribution to the bureaucratic politics literature in this context.

## References

Agranoff, Robert, and Michael McGuire. 2003. *Collaborative Public Management*. Washington, DC: Georgetown University Press.

Ashworth, Rachel, George Boyne, and Tom Entwistle. 2010. *Public Service Improvement: Theories and Evidence*, Oxford: Oxford University Press.

Bansal, Pratima, and Iain Clelland. 2004. Talking Trash: Legitimacy, Impression Management, and Unsystemic Risk in the Context of Natural Environment. *Academy of Management Journal* 47(1): 93-103.

Bajada, Christopher. 2005. Unemployment and the Underground Economy in Australia. *Applied Economics* 37(2): 177–189.

Barnett, Adrian G, Jolieke C. van der Pols, and Annette J. Dobson. 2005. Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology* 34: 215-220.

Beck, Nathaniel, and Jonathan N. Katz. 1995. What to do (and not to do) with time-series cross-section data. *American Political Science Review* 89(3): 634–647.

Blasi, Gerald J. 2002. Government Contracting and Performance Measurement in Human Services. *International Journal of Public Administration* 25: 519-38.

Boyne, George A. 2003. Sources of Public Service Improvement: A Critical Review and Research Agenda. *Journal of Public Administration Research and Theory* 13: 367-94.

Bradbury, Bruce. 1988. *Welfare Fraud, Work Incentives and Income Support for the Unemployed*. Sydney: Social Welfare Research Centre, University of New South Wales

Carpenter, Daniel P. 2001. *The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies, 1862-1928*. Princeton: Princeton University Press.

Carpenter, Daniel P. 2002. Groups, the Media, Agency Waiting Costs, and FDA Drug Approval. *American Journal of Political Science* 46 (July): 490-505.

Carpenter, Daniel P. 2004. Protection without Capture: Dynamic Product Approval by a Politically Responsive, Learning Regulator. *American Political Science Review* 98: 613-31.

Carpenter, Daniel P. 2010a. *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA*, Princeton: Princeton University Press.

Carpenter, Daniel P. 2010b. Institutional strangulation: Bureaucratic politics and financial reform in the Obama administration. *Perspectives on Politics* 8: 825-46.

Carpenter, Daniel P. and George A. Krause. 2012. Reputation and Public Administration. *Public Administration Review* 72: 26-32.

Centerlink. 2000-2002. *Annual Reports*. Canberra: Centerlink.

Centerlink. 2011. Centerlink Fraud and Compliance Media Strategy 2010-2011 Overview. Canberra: Centerlink.

Cho Yoon Jik, and Evan J. Ringquist. 2011. Managerial Trustworthiness and Organizational Outcomes. *Journal of Public Administration Research and Theory* 21 (1): 53-86

Chunn Dorothy E. and Shelley A.M. Gavigan. 2004. Welfare Law, Welfare Fraud, and the Moral Regulation of the 'Never Deserving' Poor. *Social & Legal Studies*, 13(2): 219-243.

Chunn Dorothy E. and Shelley A.M. Gavigan. 2006. From Welfare Fraud to Welfare as Fraud: The Criminalization of Poverty, In *Criminalizing Women*, eds. Gillian Balfour and Elizabeth Comack, 217-235. Black Point, Nova Scotia: Fernwood Publishing.

Cohen, Steven, and William B. Eimicke. 2008. *The Responsible Contract Manager: Protecting the Public Interest in an Outsourced World*. Washington, DC: Georgetown University Press.

Dawson, Jeremy F. 2014. Moderation in Management Research: What, Why, When, and How. *Journal of Business Psychology*, 29: 1-19.

Deephouse, David L. 1996. Does Isomorphism Legitimate? *Academy of Management Journal* 39(4): 1024-1039.

Deephouse, David L. 2000. Media Reputation as a Strategic Resource: An Integration of Mass Communication and Resource-based Theories. *Journal of Management* 26(6): 1091-1112.

Department of Human Services. 2002-2012. *Annual Reports*, Manuka ACT: Australian Government.

Driscoll, John. C. and Aart C. Kraay. 1998. Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Review of Economics and Statistics* 80: 549-560.

Dunleavy, Patrick, and Leandro Carrera. 2013. *Growing the Productivity of Government Services*. Cheltenham: Edward Elgar.

Gilad, Sharon. 2012. Regulatory Attention and Reputation, In *Executive Politics in Times of Crisis*, eds. Martin Lodge and Kei Wegrich, 157-178. Basingstoke: Palgrave Macmillan.

Gilad, Sharon and Tamar Yogev. 2012. How Reputation Regulates Regulators: illustrations from the Regulation of Retail Finance. In *Oxford Handbook of Corporate Reputation*, eds. Michael Barnett and Tim Pollock, 320-340. Oxford: Oxford University Press.

Gilad Sharon, Moshe Maor and Pazit Ben-Nun Bloom. 2013. Organizational Reputation, the Content of Public Allegations and Regulatory Communication. *Journal of Public Administration Research and Theory* DOI: 10.1093/jopart/mut041

Grover, Chris. 2005. Advertising Social Security Fraud. *Benefits* 13(3):199-205.

Halligan, John. 2007. Advocacy and Innovation in Intreragency Management: The Case of Centerlink. *Governance* 20(3): 445-467.

Halligan, John (with Jules Wills). 2008. *The Centrelink Experiment Innovation in service delivery*. Canberra: ANU E Press.

Heclo, Hugh. 1977. *A Government of Strangers*. Washington, DC: Brookings Institution.

Heimann, C. F. Larry. 1997. *Acceptable Risks: Politics, Policy and Risky Technologies*. Ann Arbor: University of Michigan Press.

Hoechle, Daniel. 2007. Robust standard errors for panel regressions with crosssectional dependence. *The Stata Journal* 7(3): 281-312.

Kahneman Daniel, and Amos Tversky. 1984. Choices, values, and frames. *American Psychologist* 39: 341–350.

Karvelas Patricia. 2008. Welfare fraud on Rudd's hit list. *The Australian* 25 January, p. 4.

Kelly, Janet M. 2005. The Dilemma of the Unsatisfied Customer in a Market Model of Public Administration. *Public Administration Review* 65: 76-84.

Krause, George A. and James W. Douglas. 2005. Institutional Design versus Reputational Effects on Bureaucratic Performance: Evidence from U.S. Government Macroeconomic and Fiscal Projections. *Journal of Public Administration Research and Theory* 15(2): 281-306.

Krause, George A. and James W. Douglas. 2006. Does Agency Competition Improve the Quality of Policy Analysis? Evidence from OMB and CBO Current Year Fiscal Projections. *Journal of Policy Analysis and Management* 25:53-74.

Krause, George A., and J. Kevin Corder. 2007. Explaining Bureaucratic Optimism: Theory and Evidence from U.S. Federal Executive Agency Macroeconomic Forecasts. *American Political Science Review* 101(1): 129-42.

Lee, Soo-Young, and Andrew B. Whitford. 2012. Assessing the Effects of Organizational Resources on Public Agency Performance: Evidence from the US Federal Government. *Journal of Public Administration Research and Theory* doi: 10.1093/jopart/mus050

Lichtenstein Sarah, Fischhoff Baruch, and Lawrence D. Phillips (1982) Calibration of probabilities: the state of the art to 1980. In *Judgment under Uncertainty: Heuristics and Biases*, eds. Daniel Kahneman, Paul Slovic, and Amos Tversky, 206-334. Cambridge: Cambridge University Press.

Linder, Jane C. 2004. *Outsourcing for Radical Change: A Bold Approach to Enterprise Transformation*. New York, NY: American Management Association.

Macmillan, William D. 2012. *Agencies and Appropriations*. Ph.D. Dissertation. University of Michigan.

Maor, Moshe. 2007. A Scientific Standard and an Agency's Legal Independence: Which of These Reputation-Protection Mechanisms Is Less Susceptible to Political Moves. *Public Administration* 85: 961-78.

Maor, Moshe. 2010. Organizational Reputation and Jurisdictional Claims: The Case of the U.S. Food and Drug Administration. *Governance* 23: 133-59.

Maor, Moshe. 2011. Organizational Reputation and the Observability of Public Warnings in 10 Pharmaceutical Markets. *Governance* 24: 557-82.

Maor, Moshe. In Press. Theorizing Bureaucratic Reputation. In Arild Wæraas and Moshe Maor (eds.) *Organizational Reputation in the Public Sector*. London: Routledge.

Maor, Moshe and Raanan Sulitzeanu-Kenan. 2013. The Effect of Salient Reputational Threats on the Pace of FDA Enforcement. *Governance* 26(1): 31-61.

Maor Moshe, Sharon Gilad and Pazit Ben-Nun Bloom. 2013. Organizational Reputation, Regulatory Talk and Strategic Silence. *Journal of Public Administration Research and Theory* 23(3): 581-608.

McKeever, Gráinne. 1999a. Fighting Fraud: A n Evaluation of the Government's Social Security Fraud Strategy, *Journal of Social Welfare and Family Law* 21(4): 357-371.

McKeever, Gráinne. 1999b. Detecting, Prosecuting and Punishing Benefit Fraud: The Social Security Administration (Fraud) Act 1997. *The Modern Law Review* 62(2): 261-270.

Moffitt, Susan L. 2010. Promoting Agency Reputation through Public Advice: Advisory Committee Use in the FDA. *The Journal of Politics* 72: 880-93.

Moore, Mark H. 1995. *Creating Public Value: Strategic Management in Government*. Cambridge, MA: Harvard University Press.

Moynihan, Donald P. 2012. Extra-Network Organizational Reputation and Blame Avoidance in Networks: The Hurricane Katrina Example. *Governance* 25(4): 567-88.

Moynihan, Donald P., and Sanjay K. Pandey. 2005. Testing How Management Matters in an Era of Government by Performance Management. *Journal of Public Administration Research and Theory* 15: 421-39.

National Audit Office (NAO). 2006. *International Benchmark of Fraud and Error in Social Security Systems*. London: Stationery Office.

Nohria, Nitin, and Ranjai Gulati. 1996. Is Slack Good or Bad for Innovation? *Academy of Management Journal* 39(5): 1245-1264.

Prenzler, Tim. 2011. Welfare Fraud in Australia: Dimensions and Issues, *Trends & Issues in Crime and Criminal Justice*. No. 421, Canberra: Australian Institute of Criminology.

Prenzler, Tim. 2012. *Responding to Welfare Fraud: The Australian Experience*, AIC Reports 119, Canberra: Australian Institute of Criminology.

Pullock, Timothy G. and Violina P. Rindova. 2003. Media Legitimation Effects in the Market for Initial Public Offerings. *Academy of Management Journal* 46(5): 631-642.

Quirk, Paul. 1980. The Food and Drug Administration. In *The Politics of Regulation*, ed. James Q. Wilson. New York: Basic Books.

Rowlingson, K., Whyley, C., Newburn, T. and Berthoud, R. (eds). 1997. *Social Security Fraud, The Role of Penalties*. DSS Research Report No. 64, London: The Stationery Office.

Russo, Michael V., and Paul A. Fouts. 1997. A Resource-Based Perspective on Corporate Environmental Performance and Profitability. *Academy of Management Journal* 40(3): 164-77.

Sandberg, Bo. 1993. Sweden: Fraud and Abuse of Welfare Systems, *Journal of Financial Crime* 5(4): 396-405.

UK Revenue & Customs and Department for Work and Pensions. 2010. *Tackling Fraud and Error in the Benefit and Tax Credits Systems*. Available online at: <http://www.dwp.gov.uk/docs/tackling-fraud-and-error.pdf> (accessed 10.1.2014).



UK Treasury & National Audit Office (NAO). 2008. *Tackling External Fraud: Good Practice Guide*. London: NAO.

Walker, Richard M., and George A. Boyne. 2006. Public Management Reform and Organizational Performance: An Empirical Assessment of the U.K. Labour Government's Public Service Improvement Strategy. *Journal of Policy Analysis and Management* 25: 371-93.

Walsh, Tamara and Greg Martson. 2010. Benefit Overpayment, Welfare Fraud and Financial Hardship in Australia, *Journal of Social Security Law* 17(2): 100-125.

Walsh, James, and Gerrardo R. Ungson. 1991. Organizational Memory. *Academy of Management Review* 16: 57-91.

Weinberger, Hadas, Dov Te'eni, and Ariel J. Frank. 2008. Ontology-based Evaluation of Organizational Memory. *Journal of the American Society for Information Science & Technology* 59: 1454-1468.

Whitford, Andrew B. 2002. Bureaucratic Discretion, Agency Structure, and Democratic Responsiveness: The Case of the United States Attorney. *Journal of Public Administration Research and Theory* 12(1): 3-27.

Whitford, Andrew B. 2003. Adapting Agencies: An Essay on Competition, Imitation, and Punishment. In *Politics, Policy, and Organizations Frontiers in the Scientific Study of Bureaucracy*, eds. George A. Krause and Kenneth J. Meier, 292-308. Ann Arbor: The University of Michigan Press.

## Appendix 1. Notable Examples of Negative and Positive Press Coverage of *Centerlink*

### *Negative Coverage*

The role of government is to look after all of its citizens. A civilized society is judged by how it treats its poor. Take the case of the single mother of one-year-old twin girls and a toddler son (*The Australian*, October 24) who had used her parenting payment for bills on the assumption that her Austudy [Department of Human Services] money would arrive as usual. She was turned away by a *Centrelink* official after the administrative glitch that caused this could not be solved. This young woman, who was doing her best to look after her family, pay her bills and study for future employment, was turned away from the agency whose fault it was that she had no money, and told to go to St Vincent de Paul for help. (*The Australian*, October 30, 2000).

Unemployed people are frequently harassed by social security staff because of the federal Government's repeated comments about "job snobs", says the Australian Council of Social Service. . . . The council is particularly critical of the *Centrelink* penalty process, which imposes fines on social security claimants if they fail to lodge the correct administrative details. . . . "These penalties are out of all proportion to the seriousness of the offence and have no rational basis," the submission says. "They are imposed because of a failure to comply with social security rules, not for instances of social security fraud." (*The Australian*, December 17, 2001).

High-end welfare cheats have been ignored while federal welfare investigators pursued small-scale debts, a government audit report has found. The Australian National Audit Office review of *Centrelink* fraud investigations has reported that auditors found 87 per cent of the cases they examined did not comply with official investigation rules. The auditors also found the *Centrelink* investigations were part of a system that favored a focus on easy targets rather than complex, serious fraud investigations. (*The Age*, October 1, 2010).

### *Positive Coverage*

Article's title: *Centrelink* fraudsters found out.

Article content: The high number of social security cheats being caught by computer technology was now more of a deterrent than the length of a jail sentence, a court was told yesterday. The Court of Appeal in Brisbane was told a computer program introduced to link the tax and social security offices led to savings of \$16 million in a trial period alone. It was now a permanent program that was catching many cheats. (*The Courier Mail*, May 6, 2010)

Investigations have been launched into statewide social security fraud by the illegal sex trade. Details of more than 700 illegal operations have been demanded from the Adult Business Association of NSW (ABA), which represents the licensed brothel industry. ABA spokesman Chris Seage said *Centrelink's* fraud investigation team, based at Lake Macquarie, had been given a list of 775 illegal brothels, massage parlors and home prostitution businesses [...] "It is good to see federal authorities are interested in pursuing

the shonks that make up the illegal brothel trade in NSW," Mr Seage said. (*The Sun Herald*, July 15, 2007).

Tasmania's crackdown on welfare fraud has paid dividends with private investigators nabbing 88% of suspected dole cheats. Twenty-four cases were investigated, yielding 21 who had made false claims. The strike rate was the second highest in Australia after northern New South Wales-Gold Coast where 92% of the 159 cases investigated revealed fraudulent claims. The investigation followed a warning by *Centrelink* that its Tasmanian "spies" were secretly videotaping suspected cheats. (*Hobart Mercury*, May 1, 2000).

Table 1: Centerlink's Outputs under Examination

Agency output	Yearly mean	Period for which data was available
Eligibility reviews	3,592,860	2000-2010
Prosecution referrals	4,227.36	2000-2010
Public campaigns	0.55	2000-2010
Data matching with government agencies	17	2000-2010
Number of formal repayment agreements	1,278,419	2000-2010
Service profile updates	835,571.7	2002-2010
Data matching with non-governmental institutions	9,889.33	2002-2010
Investigations into serious non-compliance and fraud	33,350.71	2004-2010
Job capacity assessments	292,238.2	2006-2010
Number of mercantile agents used in debt recovery	278,856.8	2006-2010

Table 2: Panel Data Regression Analyses with Driscoll-Kraay Standard Errors – standardized change in agency outputs as dependent variable

	Model 1	Model 2	Model 3	Model 4	Model 5
Valence X Lagged-output			1.508 (.425)**	1.534 (.424)**	1.586 (.408)**
Valence X Slack					.057 (.023)*
Valence		.389 (.402)	.317 (.264)	.056 (.346)	.004 (.345)
Lagged output	-.703 (.061)***	-.696 (.066)***	-.977 (.137)***	-.960 (.140)***	-.966 (.143)***
Year	.215 (.048)**	.224 (.050)**	.186 (.051)**	.149 (.021)***	.142 (.022)***
Year cube	-.004 (.0006)***	-.004 (.0006)***	-.003 (.0007)**	-.003 (.0004)***	-.003 (.0004)***
Slack ( $\frac{stand.budget_{t-1}}{Y_{ot-1}}$ )				-.006 (.011)	-.018 (.012)
Constant	-.176 (.115)	-.288 (.206)	-.136 (.148)	.007 (.110)	.041 (.112)
<i>F</i>	72.51***	118.12***	35.46***	36.29***	233.02***
R-squared	.45	.46	.48	.51	.51
<i>N</i>	80	80	80	79	79

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

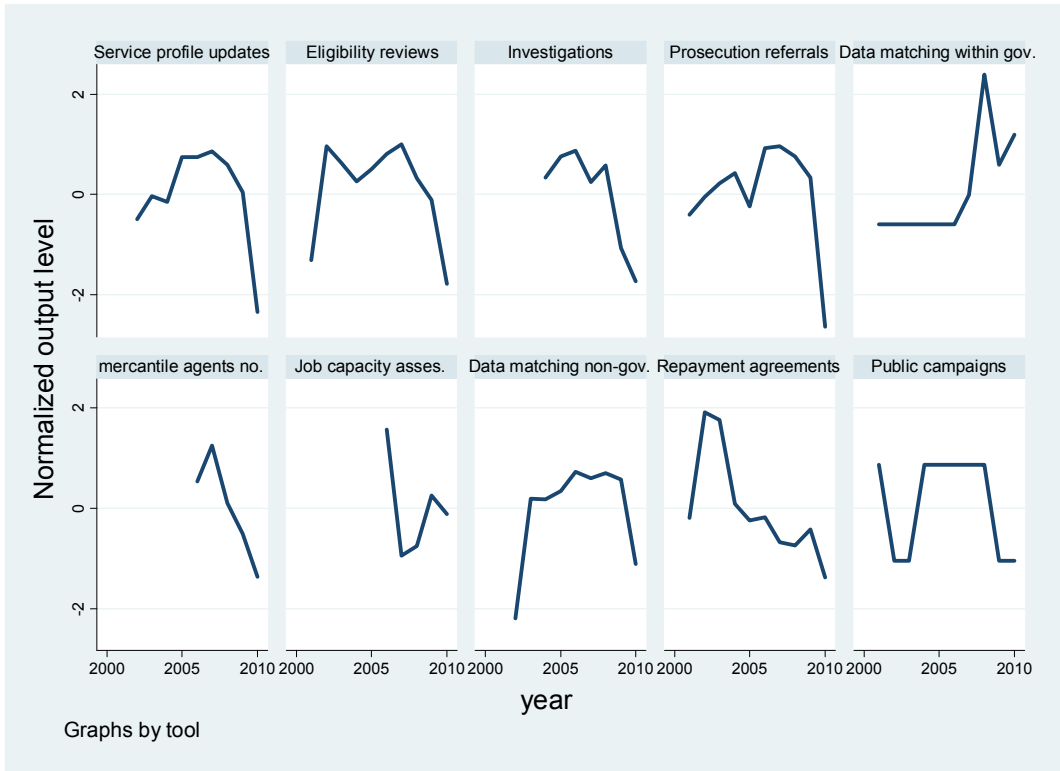


Figure 1: Temporal variation in agency outputs (measured in standard deviations)

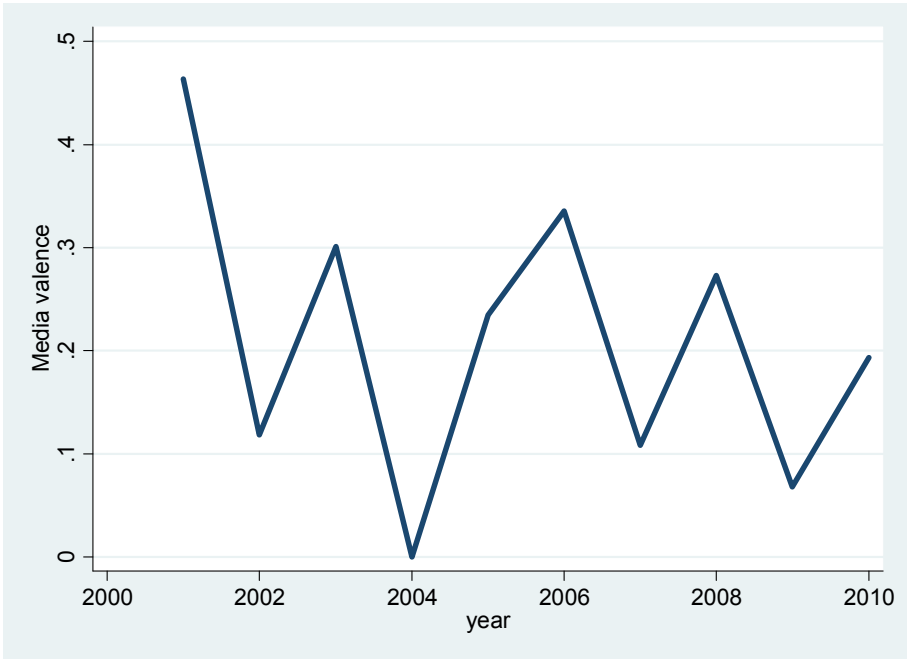


Figure 2: Temporal variation in media valence

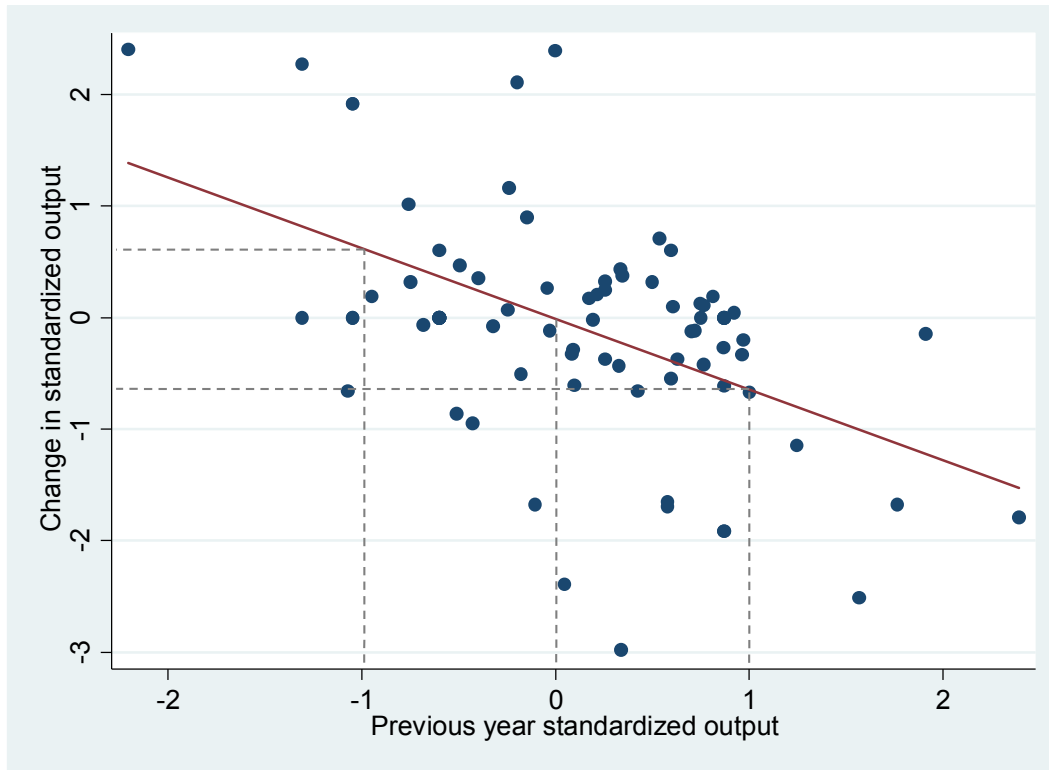


Figure 3: Estimating the magnitude of the 'regression to the mean' (RTM) effect

Mean change in standardized output as a function of one-year lagged output suggest that one standard deviation above (below) the mean is followed by average decrease (increase) of 0.634 standard deviation in the consequent year ( $p < .001$ ,  $R^2=.24$ ).



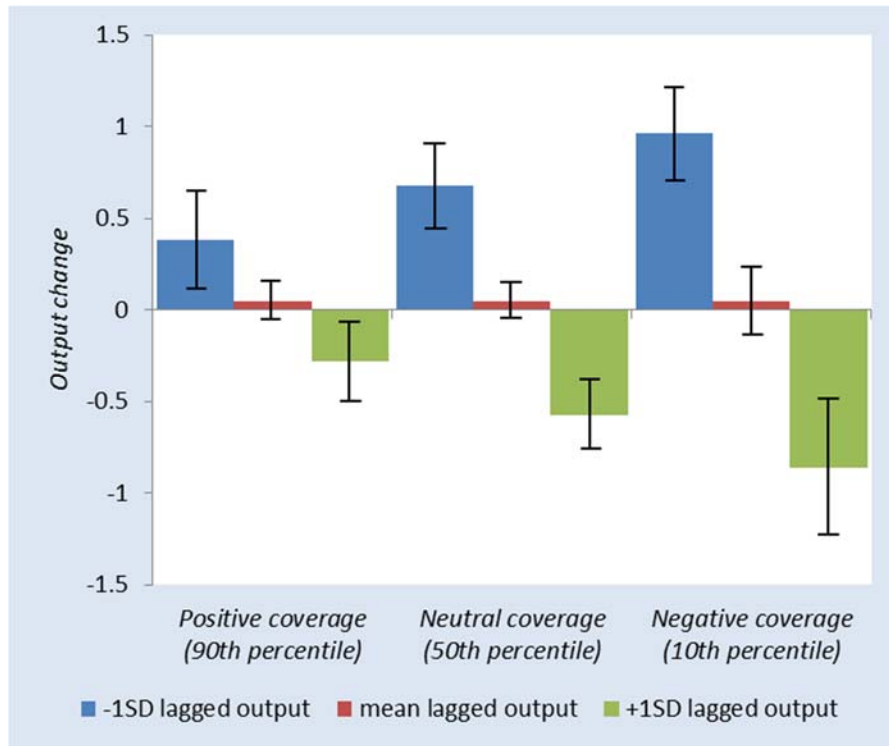


Figure 4: The relationship between lagged media valence and output change, across lagged output levels (estimates are based on model 5)

---

<sup>1</sup> Under the Human Services Legislation Amendment Act 2011, Medicare Australia and *Centerlink* services were integrated into the Department of Human Services. This major change has led to a transformation of the Department's structure and organizational culture, as well as to a change in the role of *Centerlink* (Department of Human Services 2012, 10). This, in turn, has closed the window for extending the period under examination.

<sup>2</sup> The exact search specification was as follows: "((welfare) OR (social security) OR (social benefits) OR (*Centrelink*)) AND ((fraud) OR (scam) OR (deceit) OR (cheat) OR (overpayment) OR (investigation) OR (plead guilty) OR (found guilty))".

<sup>3</sup> We are grateful to Donald Moynihan for pointing our attention to this possibility.

<sup>4</sup> An assessment of several specifications of time-trend suggests a non-linear trend, best described by a cubic relationship.

<sup>5</sup> Senior official, Department of Human Services, Interview with Moshe Maor, Canberra 31.1.2012.

<sup>6</sup> Senior official, *Centerlink*, Department of Human Services, Interview with Moshe Maor, Canberra 30.1.2012.

<sup>7</sup> Senior official, Department of Human Services, Interview with Moshe Maor, Canberra 31.1.2012.

<sup>8</sup> In the early 2000s, when most criticism of *Centerlink* has revolved around poor service delivery, mostly over its phone service center, the agency tried to build confidence in the integrity of the system. And in the mid-2000s, it advanced reputation for credibility, especially by supplying the media with covert surveillance footage of fraudulent

---

*Centerlink* customers as well as vision obtained for cash economy operations (e.g., pre-operation surveillance) (*Centerlink* 2011; Senior official, Department of Human Services, Interview with Moshe Maor, Canberra 31.1.2012).